AMERICAN JOURNAL OF PRIMATOLOGY
Official Journal of The American Society of Primatologists
ASP

Intermittent pair-housing, pair relationship qualities, and HPA activity in adult female rhesus macaques

Journal:	*American Journal of Primatology*
Manuscript ID	AJP-17-0133.R3
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Hannibal, Darcy; UC Davis, Primate Center; Cassidy, Lauren; German Primate Center Vandeleest, Jessica; University of California-Davis, California National Primate Research Center Semple, Stuart; University of Roehampton, Chun, Katie; University of California-Davis, California National Primate Research Center Barnard, Allison; University of California-Davis, California National Primate Research Center Winkler, Sasha; University of California-Davis, California National Primate Research Center McCowan, Brenda; UC Davis, SVM: Pop Health & Repro
Keywords:	pair-housing, peer interaction, cortisol, overnight separation

Intermittent pair-housing, pair relationship qualities, and HPA activity in adult female rhesus macaques

Darcy L Hannibal[1,2], Lauren C Cassidy[3], Jessica Vandeleest[1,2], Stuart Semple[4], Allison Barnard[1], Katie Chun[1], Sasha Winkler[1], Brenda McCowan[1,2]

[1]California National Primate Research Center, University of California Davis, Davis, California

[2]Department of Population Health & Reproduction, University of California Davis, Davis, California

[3]Welfare and Cognition Group, Cognitive Neuroscience Laboratory, German Primate Center, Göttingen, Germany

[4]Centre for Research in Evolutionary, Social and Interdisciplinary Anthropology, University of Roehampton, London, United Kingdom

Short title: Pair-housing and HPA activity in rhesus

Correspondence:

Darcy Hannibal, PhD

California National Primate Research Center

University of California

One Shields Ave

Davis, CA 95616

Phone: 530-752-1586

Email: dlhannibal@ucdavis.edu

ABSTRACT

Laboratory rhesus macaques are often housed in pairs and may be temporarily or permanently separated for research, health, or management reasons. While both long-term social separations and introductions can stimulate a stress response that impacts inflammation and immune function, the effects of short-term overnight separations and whether qualities of the pair relationship mediate these effects are unknown. In this study, we investigated the effects of overnight separations on the urinary cortisol concentration of 20 differentially paired adult female rhesus macaques (*Macaca mulatta)* at the California National Primate Research Center. These females were initially kept in either continuous (no overnight separation) or intermittent (with overnight separation) pair-housing and then switched to the alternate pair-housing condition part way through the study. Each study subject was observed for five weeks, during which we collected measures of affiliative, aggressive, anxious, abnormal, and activity-state behaviors in both pair-housing conditions. Additionally, up to three urine samples were collected from each subject per week and assayed for urinary free cortisol and creatinine. Lastly, the behavioral observer scored each pair on four relationship quality attributes ("Anxious," "Tense," "Well-meshed," and "Friendly") using a seven-point scale. Data were analyzed using a generalized linear model with gamma distribution and an information theoretic approach to determine the best model set. An interaction between the intermittent pairing condition and tense pair adjective rating was in the top 3 models of the best model set. Dominance and rates of affiliation were also important for explaining urinary cortisol variation. Our results suggest that to prevent significant changes in HPA-axis activation in rhesus macaque females, which could have unintended effects on research outcomes, pairs with "Tense" relationships and overnight separations preventing tactile contact should be avoided.

Keywords: pair-housing; overnight separation; peer interaction; cortisol

INTRODUCTION

It is well established that social environments, compared to solitary housing, significantly improve captive non-human primate (NHP) welfare and health (Olsson & Westlund, 2007). For example, single-housing has been associated with physiological changes, such as higher blood pressure (Coelho, Carey, & Shade, 1991) and immunosuppression (Lilly, Mehlman, & Higley, 1999), that increase the risk of acquiring pathological health conditions (e.g., cardiovascular disease or infection). Furthermore, studies in laboratory rodents have demonstrated that environments lacking complexity, such as limited cage features and insufficient outlets for expressing species adaptations, can have deleterious effects on biomedical research results (e.g., Richter et al., 2011). Consequently, regulatory pressure has increased on research facilities to socially house NHPs (Hannibal, Bliss-Moreau, Vandeleest, McCowan, & Capitanio, 2017). Although social housing is the expected and enforced norm, laboratory NHPs may experience extended periods of social separation due to colony or study protocols. For example, pair-mates may be separated to prevent a social partner from picking at and removing surgical sutures, confirm diarrhea or menses after overnight separation, or collect overnight urine or fecal samples. The effects of these separations on the welfare, physiology, and health of laboratory NHPs are not well understood. In this paper, we investigate the effects of daily, overnight separations of paired adult female rhesus macaques (*Macaca mulatta*) on urinary cortisol, a hormonal measure that is sensitive to environmental changes and reflects physiological states that may impact research outcomes.

Among all research facilities in the United States, laboratory NHPs are primarily housed in social groups (61.51%), less often in pairs (22.84%), or singly-housed (15.65%) (Bennett, 2016).

Pair-housing, the cohousing of two individuals by connected adjacent cages, has been developed and refined to maximize social contact for laboratory NHPs in a manner compatible with many research objectives (Baker, Crockett, et al., 2012). Single-housing facilitates specific research objectives, but maintains individuals in separate cages. Although this allows auditory, visual, and olfactory contact with conspecifics, tactile contact is restricted to varying degrees depending on whether the separating door is solid metal, bars, grate, or mesh (Baker, Bloomsmith, et al., 2014; Bennett, 2016). Single-housing, however, is prohibited by regulations, unless justified by clinical or behavioral findings that require pair separation or research needs that have been reviewed and approved by the institutional oversight office (United States Department of Agriculture, 2013). Modified forms of pair-housing are often used to accommodate research or management needs. Intermittent pair-housing involves temporary daily or weekly separations that last 12 or more hours, including overnight (Baker, 2016; Capitanio, Blozis, Snarr, Steward, & McCowan, 2017). In contrast, continuous pair-housing allows complete visual and physical access to a pair-mate, with infrequent and brief separations.

Several studies have demonstrated welfare improvements for NHPs that are pair-housed as compared to those that are singly-housed. For example, pair-housing has been associated with improved behavioral welfare indices, including reduced levels of abnormal and anxiety-related behaviors (e.g., Baker, Bloomsmith, et al., 2012; Gottlieb, Maier, & Coleman, 2015), enhanced repertoires of species-specific behaviors (e.g., Baker, Bloomsmith, et al., 2014), and decreased self-injurious behavior (SIB) (e.g., Rommeck, Anderson, Heagerty, Cameron, & McCowan, 2009; Weed et al., 2003). Another study found that pair-housed NHPs had better immune function than single-housed NHPs (Schapiro, Nehete, Perlman, & Sastry, 2000). While the benefits of pair-housing are now well established, pairing laboratory macaques with compatible

companions is challenging and requires knowledge of and experience with species-specific social behavior (Truelove, Martin, Perlman, Wood, & Bloomsmith, 2017). Thus, research on laboratory macaque pair-housing has shifted focus to refining pairing practices to improve partner compatibility, welfare, and pairing success (e.g., Capitanio et al., 2017; Pomerantz & Baker, 2017; Truelove et al., 2017). Relatively little progress has been made, however, to improve our understanding of how frequent changes to pair-housing affect NHP physiology, despite the implications for biomedical research (reviewed in Hannibal et al., 2017).

Captive NHPs tend to have better welfare measures when they are able to express key species-specific behaviors (Lutz & Novak, 2005). Although most primate species spend a significant amount of their activity budget engaged in social behavior (Dunbar, 1991), captive pair-housed NHPs spend even more time doing so (Crockett, Bowers, Bowden, & Sackett, 1994), likely due to a limited repertoire of other activities. For both wild and captive NHPs, the longest bouts of affiliation occur when they are huddled together overnight (Anderson, 1998; Eaton, Kelley, Axthelm, Iliff-Sizemore, & Shiigi, 1994). Furthermore, NHPs actively prefer the proximity of a social partner even when there are costs associated with that choice. For example, adult rhesus macaques chose to remain in the same cage as their social companions despite tradeoffs in available space (Basile, Hampton, Chaudry, & Murray, 2007). Also, captive tufted capuchin monkeys (*Cebus apella*) often chose their companions over food, even several hours after food deprivation (Dettmer & Fragaszy, 2000). Lastly, access to social partners buffers physiological stress during stressful procedures in captivity (Hennessy, Kaiser, & Sachser, 2009; Kikusui, Winslow, & Mori, 2006; Truelove et al., 2017), such as witnessing the anesthesia of another animal in the room (Gilbert & Baker, 2011).

In contrast, separations from conspecifics can negatively impact NHP behavior and physiology. Physiological disruptions associated with permanent social group removal are "substantial" and take about 3-months to return to baseline, thus a 3-month conditioning period is recommended when previously outdoor housed NHPs are moved into indoor research settings (reviewed in Capitanio, Kyes, & Fairbanks, 2006). Temporary separations from social contact for greater than 10 hours to several days, are also known to increase negative indices of welfare in captive NHPs. For example, adolescent rhesus macaques displayed higher levels of abnormal and depressive behaviors in response to a 4-day social separation, increasing further after repeated separations (Mineka, Suomi, & DeLizio, 1981). Also, an 11-hour period of social isolation in Wied's black tufted-ear marmoset monkeys (*Callithrix kuhli*) was associated with increased urinary cortisol concentration (Smith & French, 1997).

While the implementation of intermittent pair-housing varies among facilities, all cases involve at least some overnight separation, as previously mentioned (Baker, 2016; Capitanio et al., 2017; Roberts & Platt, 2005; Rommeck, Capitanio, Strand, & McCowan, 2011; Tardif, Coleman, Hobbs, & Lutz, 2013). Continuously paired animals still experience short daytime separations for sample collection, health checks, and husbandry procedures, but spend more than half of every day together, with the exception of serious, albeit rare, health issues. At the California National Primate Research Center (CNPRC), intermittently housed monkeys are separated from about 14:00 (just prior to the afternoon feeding) until 08:00 (after the morning feeding) the following day, providing a maximum of 6 hours of daily socialization and physical contact. These separations remove the opportunity for these individuals to receive the benefits of overnight social contact (Eaton et al., 1994; Kikusui et al., 2006). Therefore, the welfare of

intermittently pair-housed NHPs needs to be characterized by incorporating indices of welfare that can capture the lasting effects of overnight separations.

Physiological indices of welfare, specifically the measurement of hypothalamic-pituitary-adrenal (HPA) axis activity, can provide insight into the impacts of overnight social separation. The main output of the HPA axis is cortisol, a glucocorticoid that can influence a variety of physiological systems, especially those involved in stress response and immune functioning (Sapolsky, Romero, & Munck, 2000). Depending on the biological source, elevated HPA axis activity can be detected several minutes (blood), hours (urine), days (feces), or months (hair) after a stressor has occurred (Novak, Hamel, Kelly, Dettmer, & Meyer, 2013). Activity of the HPA-axis is known to be highly sensitive to environmental influences (e.g., temperature, stress) (Herman et al., 2003; Vandeleest, Blozis, Mendoza, & Capitanio, 2013) including the social environment (Mendoza, Capitanio, & Mason, 2001). Social isolation and unstable social relationships can lead to elevated cortisol levels and, when chronic, can eventually lead to altered regulation of the HPA axis (Capitanio, Mendoza, Lerche, & Mason, 1998; Dettmer, Novak, Meyer, & Suomi, 2014). For example, wild male olive baboons (*Papio anubis)* that were about to lose rank had higher cortisol levels than similarly ranked males that were about to gain rank (Sapolsky, 1992). On the other hand, higher rates of positive social interactions, like grooming, have been associated with lower fecal cortisol concentrations in Barbary macaques (*Macaca slyvanus*) (Shutt, MacLarnon, Heistermann, & Semple, 2007) and with lower hair cortisol concentrations in rhesus macaques (Wooddell et al., 2017). Relative cortisol levels, thus, are only useful when informed by the context (climate, activity, rank relationships, and other social and environmental variables) and perturbations associated with changes in levels.

Research on the impact of social housing (pair- vs single-housing) on cortisol levels has yielded mixed results. Although some previous studies found no differences in serum cortisol concentrations between single- and pair-housed macaques (e.g., Baker, Bloomsmith, et al., 2012; Gust, Gordon, Brodie, & McClure, 1994; Schapiro, Bloomsmith, Kessel, & Shively, 1993), others have found higher cortisol levels in singly-housed animals (Doyle, Baker, & Cox, 2008). These studies, however, vary in a couple of potentially important ways. First, they differ in the sampling matrix used to measure cortisol levels. All of the studies failing to find a relationship between cortisol and pairing status measured serum cortisol levels, whereas the Doyle et al. (2008) study measured fecal cortisol levels. These sampling matrices reflect HPA-axis activation on a scale of minutes (serum) to days (feces) which may have impacted the measured relationships. Secondly, these studies varied in whether, or the degree to which, they pre-selected potential pair-mates based on criteria that tend to maximize compatibility (e.g., body weight disparity). Since positive and negative social interactions can alter HPA axis activation, the qualities of the pair relationship may be critical to the ability to detect differences in cortisol levels. Overall, the consequences of manipulating a NHPs' social environment (e.g., switching between pair-housing conditions) on their behavior and physiological functioning remain largely unknown (Hamel et al., 2017; reviewed in Hannibal et al., 2017). Pair-mate compatibility may alter the magnitude of the stress response to pair separations and reunions. Therefore, investigating the pair relationship could uncover behavioral compatibility metrics that are likely to facilitate less stressful separations and reunions. It is unlikely that there is a single metric of pair compatibility, but converging evidence from more than one behavioral or physiological metric would allow managers to use the metrics they have access to and that have predictive power.

In this study, we investigate whether changes in intermittent versus continuous pair-housing condition of adult female rhesus macaques impacts the HPA axis as measured by urinary cortisol concentrations. We further explore the impact of pair relationship quality and whether it modulates the effect of housing condition, while controlling for other aspects of the social environment, such as dominance status and affiliation rates. For NHPs adapted for a rich social life, long periods of social isolation have the potential to produce physiological variability with implications for the external validity of biomedical research conducted with such animals (Hannibal et al., 2017). If overnight separation is associated with substantial changes in HPA axis activity, then modifications of this practice should be considered for the benefit of both animal welfare and research.

METHODS

This research was conducted from March to May 2015 at the California National Primate Research Center (CNPRC) in Davis, California. Animal care and research protocols for this study were approved by the Institutional Animal Care and Use Committee at the University of California Davis. This research was conducted in accordance with United States federal regulations and adhered to the American Society of Primatologist Principals for the Ethical Treatment of Animals.

Subjects

In order to limit physiological variability of the study sample as much as possible, subject selection criteria included: (a) only females due to sex differences in physiology and the fact that most adults in the indoor colony are female; (b) a minimum three months indoors and in their pair-housing condition, without repeated incidents of serious physical aggression and wounding; (c) no history of conception during the past breeding season, (d) reared in an outdoor social

group, and (e) between 4 to 11 years old, (criteria based on findings and recommendations by Capitanio et al., 2006; Cavigelli & Caruso, 2015; Reeder & Kramer, 2005). Subjects were enrolled as pairs as much as possible to avoid pair separations for other colony or project needs not related to this study. Random selection and assignment of animals was not possible because the purpose of the study was to understand impact of indoor pairing practices on physiology and the pool of animals that fit our selection criteria was very small.

The study began with 24 adult female rhesus macaques. Due to our subject criteria, 2 females were enrolled in the study while their pair-mates were not. To maintain consistency in behavioral data collection and conduct pair-adjective ratings, these data were collected on both pair-mates for all subjects, but data from the 2 non-study pair-mates of subjects was not included in individual level analyses. Two study subjects, who were paired together, were dropped during the study due to intra-pair conflict and another two were dropped from analyses due to poor or insufficient urinary samples, leaving 20 subjects. Subjects were ages 4.9 to 10.9 years (mean=6.7, SD=1.8), confirmed non-pregnant by ultrasound, and were not observed to have a consistent pattern of menstrual synchronization within pairing groups (i.e., females cycled at different times throughout the study). All subjects were born and raised in outdoor large (0.2-hectare outdoor enclosures containing up to 180 NHPs) or small (43.7 m^2 outdoor enclosures containing up to 30 NHPs) social groups comprised of all age and sex classes at the CNPRC for at least the first 2.5 years of life. Subjects selected for the study had been relocated for management reasons to indoor housing at least four months prior to the study (mean=20.7, SD=20.0). All subjects had been housed successfully (without persistent agonism or wounding) with another female in their baseline condition (intermittent or continuous) for at least three

months prior to the study (mean=11.0, SD=6.4). The baseline pairing condition was intermittent for 9 subjects and continuous for 11 subjects.

Housing and Pairing

Animal housing consisted of pairs of stainless steel cages (floor space 0.4 m^2, height 0.8 m). The cages of paired animals were joined by an opening (approximately 30 cm by 30 cm) with a sliding solid stainless-steel partition that prevented physical contact. Per management practice, intermittently paired animals were separated by the partition prior to afternoon feeding (approximately 14:00) until after morning feeding (approximately 08:00) the following day. Therefore, intermittent pairs had about six hours of co-housing each day. Conversely, continuously paired animals were co-housed for at least 18 hours daily and were always together overnight. All socially housed animals in the colony, regardless of housing condition, experience occasional separations for minutes, hours, or even days for sample collections, veterinary exams or treatments, and husbandry procedures. However, unless intermittently-housed, the majority of their 24-hour days are spent in social contact. For the purposes of urine sample collection and feeding regime consistency across the experimental groups, continuous pairs were separated during each feeding time (two bouts) for about an hour in the morning, and one to three hours in the afternoon (cumulative maximum of four hours per day). Afternoon feeding time coincided with urine sample collection for all pairs, to ensure correct identification and prevent cross-contamination of samples. Continuous pairs were re-paired immediately after an adequate sample was obtained or as soon as the 4-hour mark was reached. Intermittent pairs remained separated overnight, consistent with the colony management protocol for this housing category. The short separations of continuously housed subjects for sample collection is not of long enough duration to be considered intermittent because they were only long enough to obtain

samples and they were not separated overnight. Subjects were fed a standard monkey chow diet and a forage mixture of rice, split peas, and oats twice daily by animal care staff, with fresh water available *ad libitum*. Regular facility enrichment (e.g., mirror, chew toy, forage board, metal perch, puzzle feeders) was provided to each subject according to CNPRC standard operating procedures (SOPs) throughout the study.

Experimental Design

To compare the behavior and urinary cortisol concentration of continuously (C) versus intermittently (I) pair-housed female rhesus macaques, subjects were assigned to one of two experimental groups (i.e., CI or IC) based on their pairing condition at the beginning of the study (i.e., initial pairing condition; variable definitions are listed in Table 1). Pairs were in their initial pairing condition for two weeks (i.e., initial project phase), and then switched to their experimental condition for three weeks (i.e., experimental project phase) (Fig 1). Because it was not possible to complete data and sample collection on all subjects in one five-week study period, subjects were studied in two cohorts, balanced by experimental group so that there were about equal numbers of CI and IC subjects in each cohort. The first cohort was studied March 23 to April 24, 2015 and the second cohort was studied April 27 to May 29, 2015. All data collection occurred on weekdays (initial project phase: 9-10 days; experimental project phase: 14-15 days).

Behavioral Data Collection

Two eight-minute focal observations were conducted on each pair per observation day, between 11:15 and 13:45 hours in a randomized order. Affiliative, agonistic, status, activity, self-directed, and abnormal behaviors (see Table 1 for variables comprised of these behaviors) were recorded using the HanDBase application (DDH Software, Wellington, Florida, USA) on an

Android tablet. Observations were conducted solely by co-author L. C. Cassidy, who was previously trained as a CNPRC behavioral management staff member and reliable on all observation ethograms. For each pair, 18-20 observations were conducted during the initial pairing condition and 26-28 observations were conducted during the experimental pairing condition. Behaviors were recorded using one-zero sampling with 20-second sample intervals, except for self-directed behaviors which were recorded using all occurrences event sampling. For each observation day, proportions and frequencies were calculated for behaviors recorded with one-zero sampling and event sampling, respectively.

[INSERT TABLE 1 HERE]

Pair Relationship Adjective Ratings

Four pair rating adjectives, "Anxious," "Tense," "Well-meshed," and "Friendly" (see Table 1, Pair Rating variables), defined by co-author K. Chun, were used to evaluate the relationship of each pair on a seven-point scale. Adjective ratings allow observers to integrate multi-modal information about animals across time and experiences, and can be scientifically tested for reliability and validity (Meagher, 2009). Dyad ratings have been used to assess social interactions between amygdala lesioned vs. control animals (Emery et al., 2001). Like personality ratings, these adjectives likely remain relatively constant across different contexts (Capitanio, 1999). It was not our aim to use adjective ratings to assess possible changes to pair relationships between the initial and experimental project phases. Rather, we incorporated them to have an overall assessment of qualities of the pair relationship, irrespective of Project Phase, to assess whether this had an impact on potential changes in physiological responses to the experiment. Pair adjective ratings for the current study were assessed based on the behavioral observer's (co-author L.C. Cassidy) direct experience with the subjects over the study period.

Ratings were conducted one to two days after the data collection period for each cohort concluded, and again nine days later to assess intra-observer reliability using Krippendorf's α for interval metrics (Anxious α=0.92; Tense α=0.88; Well-meshed α=0.93; Friendly α=0.84) (Hayes & Krippendorff, 2007). The mean of the two observations for each reliable pair adjective rating was used in analyses.

Urine Sample Collection

Throughout the five-week study period, urine samples were collected from each subject between the hours of 14:00 and 17:00 each weekday until up to three urine samples over 3mL in volume (considered an adequate sample) were collected for that week. The limited and consistent collection period allowed minimal separation of the pairs and reduced variation in cortisol levels due to diurnal variation in primates (Novak et al., 2013). Urine was collected from clean stainless steel cage pans placed underneath each subject's cage. The pans were periodically checked for urine and a maximum volume of 45mL was transferred into a 50mL polypropylene vial. 351 samples were collected from 22 animals. On the day of collection, urine samples were stored at room temperature until 18:00. Lastly, the samples were centrifuged at 2500 RPM for five minutes to remove impurities (e.g., food particles), and the supernatant transferred to 5 mL polypropylene vials and stored at -80° C until assay.

Cortisol Assays

Urinary free cortisol (Co) was measured using a quantitative competitive immunoassay and direct chemiluminescent technology developed and conducted by the CNPRC Primate Assay Laboratory Core. A total of 313 urine samples were assayed in duplicate for this study. Analytical sensitivity of the cortisol immunoassays was 2 ng Co/mL. Inter-assay coefficient of variation (CV) was 3.1% and intra-assay CV was 1.6%. Creatinine (Cr) was measured by a

colorimetric assay to control for variations in subject body weight, urine output, and water content in each sample (Novak et al., 2013). Analytical sensitivity was 0.05 mg Cr/mL, inter-assay CV was 1.2%, and intra-assay CV was 0.5% for the creatinine assays. Urine sample concentration was normalized by dividing the cortisol concentration by the corresponding creatinine concentration. Urine samples with a creatinine concentration of 0.20 mg Cr/mL and below were excluded (n=55) from our analysis as they could have resulted in falsely elevated normalized cortisol concentrations. Of 258 urine samples that met our analysis criteria, urinary cortisol per creatinine ranged from 32.32 ng Co/mg Cr to 1617.73 ng Co/mg Cr (mean=362.29, SD=283.34 ng Co/mg Cr).

Data Analysis

Data were analyzed in Stata 14.1 using a generalized linear mixed model (GLMM) for a gamma distribution (meglm command) (Hardin & Hilbe, 2007). Both subject and pair identity were considered as potential random effects. An information theoretic (IT) approach was used to evaluate models based on goodness-of-fit, sample-size-corrected Akaike Information Criterion (AICc) scores, and differences in AICc scores (ΔAICc) following methods described by Burnham and Anderson (2002) and Burnham, Anderson, and Huyvaert (2011). We included variables in the models that the literature indicates may have an impact on HPA axis activity (e.g., menses and activity) or pair compatibility (e.g., affiliation and agonistic behavior), as well as the specific variables (Current Condition and pair adjective ratings) of interest to our research questions (see Table 1 for a list of all variables). The random effects were evaluated before considering models with fixed effects and only subject ID alone was retained as a random effect. Collinear variables were not used in the same model and among collinear variables, the variable with the lowest AICc score was retained for further model comparison.

Supplementary Table 1 contains a list of all models tested and the reasons these models were rejected from consideration. We considered all models that had both a model chi-square indicating a minimally good model and an AICc score less than the random effects only model, which indicates whether a model is better than a model with no predictors. Models violating the principal of parsimony were excluded (Burnham & Anderson, 2002). Model likelihoods, Akaike weights, and evidence ratios, which measure the strength of the evidence for these models, were calculated for a candidate set of models with a ΔAICc $\leq$7.0 (Burnham & Anderson, 2002; Burnham et al., 2011; Grueber, Nakagawa, Laws, & Jamieson, 2011; Symonds & Moussalli, 2011). From this candidate model set, a best model set was then selected based on evidence ratios $\leq$ 10 and weights were then renormalized (Burnham & Anderson, 2001, 2002). The Akaike weights for the best model set were used to calculate variate weights by summing the model weights for each variate across all models in which it was included (Burnham & Anderson, 2002). Variate weights measure the relative importance of each variate for understanding the outcome, with 1 indicating it has the highest possible certainty of being important. Marginal effects (margins command) and plots (plot command) were produced from the top model for predictors of interest (Hardin & Hilbe, 2007).

RESULTS

Of the models predicting urinary cortisol levels in our study animals, nine had at least some support with ΔAIC $\leq$ 7 (Burnham & Anderson, 2002) and were further examined as the candidate set of models (see Supplementary Table 2). From this set of candidate models, a set of best models with evidence ratios < 10 were selected and the model weights renormalized (see Table 2) (Burnham & Anderson, 2001). The Akaike weight of the best model (M1) was 0.481; therefore, there was not strong enough evidence to rely on this as the single best model and

information from other models in the best model set should also be considered. All models in the best model set contained main effects for Dominant (descriptives of categories: dominant=50.0%, subordinate=50.0%), Affiliation (descriptives of percent of observation period: mean=36.7, min=0.0, max=100.0, sd=30.5), and Current Condition (descriptives of categories: continuous=47.9%, intermittent=52.1%). In addition, the three models with the highest weight (M1, M2, and M3, w=0.929) also contained a main effect for Tense (descriptives of score: mean=2.7, min=1.5, max=5, sd=1.4) and an interaction between Tense and Current Condition. The cumulative weight of Models 1 and 2 was 0.81 and the only differences between models 1 and 2 were the main effects of Experimental Group (seen in model 1, but not 2) (descriptives of categories: CI=55.0%, IC=45.0%) and Project Phase (seen in model 2, but not 1) (descriptives of categories: initial=42.5%, 57.5%). Model 4 included Project Phase, which also occurred in model 2, as well as Total Pairing Time (descriptives of months: mean=11.5, min=3.6, max=24.8, sd=6.4) and Inactivity (descriptives of percent of observation period: mean=38.6, min=0.0, max=100.0, sd=24.3), which occurred in no other models in the best model set.

[INSERT TABLE 2 HERE]

The predictors in the best models are listed by order of importance based on their corresponding variate weights (the sum of the model weights for the models containing variate *j* and denoted as $w_+(j)$) in Table 3. All models included the main effects of Dominant, Affiliation, and Current Condition and thus all had $w_+(j)=1$. Tense and the interaction of Current Condition and Tense occurred in the top three models and had $w_+(j)=0.93$. Experimental Group ($w_+(j)=0.48$) only occurred in Model 1, Project Phase ($w_+(j)=0.40$) only occurred in Model 2 and

Model 4, and both Total Pairing Time and Inactivity (both had $w_{+}(j)$=0.07) only occurred in Model 4.

[INSERT TABLE 3 HERE]

The results of the best model (Model 1) are presented in Table 4. Dominant animals had urinary cortisol levels that were nearly half of those in subordinate animals (β=-0.497) (see Fig 2a). An increase in affiliation by ten percentage points was associated with 0.029 times lower (about three percent lower) cortisol levels (β=-0.003) (See Fig 2b). Although significant, the main effect of Current Condition was relatively small with an increase in urinary cortisol of about 0.12 times when Tense was at the mean value (2.73) for the sample (β=-0.604, exponentiated β =0.547). The main effect of Tense was not significant. The interaction of Current Condition (intermittent) and Tense was significant, but in the continuous condition for Current Condition, urinary cortisol levels stayed relatively low at all Tense ratings, while in the intermittent condition, urinary cortisol levels increased by 1.23 times as pair Tense rating increased (β=0.262) (See Fig 3).

[INSERT TABLE 4 HERE]

[INSERT FIG 1 HERE]

[INSERT FIG 2 HERE]

[INSERT FIG 3 HERE]

DISCUSSION

Our study aimed to explore the impact of temporary overnight separations due to intermittent pair-housing on adult female rhesus macaques' HPA axis activity, indexed through urinary cortisol concentrations. In addition to stress, other factors such as activity level and ambient temperature can affect cortisol secretion. For this reason, it is not possible to identify a

normal cortisol range for a species, population, or even an individual that is indicative of distress, eustress, or lack of stress. Our results showed that overnight separations were associated with higher concentrations of urinary cortisol, but that this association was dependent on key characteristics of the pair relationship and occurred even when accounting for other variables known to influence the production of cortisol. Most interestingly, pairs rated as having more tense relationships had higher urinary cortisol levels, but only when they were intermittently paired. Additionally, dominance status and greater rates of affiliation were associated with lower urinary cortisol.

Intermittent pairing, relationship quality, and urinary cortisol

Females that had more tense relationships with their partners had urinary cortisol levels 1.5-3 times higher, depending on the tense rating and variability, when intermittently paired than when continuously paired. A high pair rating for Tense may indicate that the relationship is tenuous and overnight separation may be introducing uncertainty in re-establishing the relationship when reunited. Uncertainty in dominance relationships has been associated with higher levels of pro-inflammatory proteins and greater risk for diarrhea for rhesus macaques living in large outdoor social groups (Vandeleest et al., 2016). This measure of uncertainty may indicate that a poor fit in the social group is associated with poorer health outcomes. Although cortisol is not a direct measure of health (cortisol values can have implications for health, but can also vary for reasons that have nothing to do with health outcomes), it is often used as a biomarker for increased health risk due to its responsiveness to stressors and role in regulating immune function (Sapolsky et al., 2000). Our findings are also consistent with a study in wild hamadryas baboons (*Papio hamadryas ursinus)* where relationship quality (measured as a

grooming diversity index) was related to HPA axis activity (Crockford, Wittig, Whitten, Seyfarth, & Cheney, 2008).

Although we did not find a difference in urinary cortisol concentration between intermittent and continuous housing conditions among pairs who did not have a Tense relationship, we caution against interpreting this as evidence that overnight separation does not cause distress or impact research outcomes. There may be differences among less Tense pairs that could not be detected in the sample used in this study. We suspect that a larger sample would find an effect, albeit a smaller one than that seen in Tense pairs.

When making decisions about pairing laboratory NHPs, behavioral and facility managers often have a limited number of potential partners to select from and attempt pair introductions depending on factors such as indoor population size, study needs, and breeding needs. While some of these potential pairs do not remain paired past the introduction period due to conflict, those that do and become established pairs usually remain paired until there is a management reason to separate them. Therefore, it is not surprising that we found variation in relationship quality in our sample. Since it is likely that other laboratory NHP facilities have pair-housed populations with similar variation in the quality of pair relationships, our results suggest that when pairs show signs of being tolerably, but not ideally, compatible (e.g., absence of physical affiliative social interaction or sitting in proximity to one another), it is best to avoid overnight separations to prevent uncertainty at reintroduction and unusual disruptions in their physiology.

In our study, continuous pair-housing provided near constant social interaction and was associated with reduced HPA activation, regardless of pair quality. However, continuous housing is not compatible with some research objectives. For example, biological sample collection (e.g., feces, urine) often requires that pair-mates are separated for some time (e.g., overnight) to

acquire samples from the correct subject. In these situations, providing some contact could limit unintended social consequences or changes to physiology. For example, when overnight separation is necessary, a grate or bar (as opposed to solid) divider that allows some visual and tactile access to pair-mates may be preferable.

To our knowledge, our study is the first to use subjective ratings to assess pair compatibility after pair introduction. Subjective rating assessment is an underutilized tool within the field of captive NHP welfare, despite the potential utility of animal caretaker knowledge. Furthermore, ratings are less time intensive than formal behavioral observations, are non-invasive unlike some physiological measurement techniques, and are scientifically valid when appropriately designed (Meagher, 2009). The interaction between housing condition (during intermittent pairing) and the quality of the pair relationship provides further support that ratings are associated with biological phenomena, in this case, changes in HPA activity. Interestingly, the IC pair we excluded from our analyses due to aggression and subsequent separation during the continuous pairing phase was rated as having a very tense relationship. These females previously knew each other from a large outdoor social group, but familiarity does not always translate to compatibility. Therefore, pair adjective ratings such as high Tense scores may act as useful guidelines for re-evaluating pair compatibility and guiding social management decisions.

Dominance rank and urinary cortisol concentration

Our study found that urinary cortisol was lower in dominant females than in subordinate females. Therefore, including dominance status in the model was important for interpreting the association between housing condition and HPA axis activity. Primate studies of cortisol usually find an effect of social rank, but the direction of the effect is not consistent across studies (e.g., Abbott et al., 2003; Muller & Wrangham, 2004; Shively, 1998). However, it is important to note

that social status or high cortisol values alone cannot be interpreted as distressing. Generally, if an individual is maintaining a healthy weight and social injuries are rare and minor, there is no reason to interpret their situation as deleterious.

Affiliation and urinary cortisol concentration

Greater frequency of affiliative behavior with a pair-mate was also associated with lower urinary cortisol in our study. This is consistent with previous findings that affiliative social partners dampen behavioral and physiological stress responses (Hennessy et al., 2009; Kikusui et al., 2006; Wooddell et al., 2017), but like dominance status, this cannot be used to make direct inferences about stress levels in this study sample. Because affiliation was an important predictor of urinary cortisol levels, including it in our multivariate analysis was necessary to understand any association with housing condition.

Pair compatibility criteria during pair introduction vary by facility (Baker, Coleman, Bloomsmith, McCowan, & Truelove, 2014), but generally the absence of deleterious aggression, wounding, food monopolization, and presence of status signals establishing dominance are prioritized over rates of affiliative behaviors between pair-mates. In NHPs, affiliative behaviors reinforce social bonds and frequent affiliation between individuals indicates the strength of the relationship (Silk, Altmann, & Alberts, 2006). The absence or reduced frequency of affiliation may not cause external injury, but may indicate the pair is not experiencing the full benefits of social housing.

Summary

Overall, our results emphasize that changes to the pair-housing arrangement, in combination with aspects of a pair's social relationship, can modulate urinary cortisol concentration in pair-housed adult female rhesus macaques. Importantly, although intermittent pair-housing provides

superior welfare over single-housing, our results indicate that it may be associated with increased HPA-axis activity when the relationship between the two pair-mates is tense. Our findings support the importance of assessing compatibility between pair-mates beyond the current minimum criteria of the absence of serious injury and repeated fighting. We also caution against interpreting the lack of an effect found for Tense pairs in this study as evidence that overnight separations do not have an impact on welfare or research as this may have been detected if a larger sample was possible.

We propose a continuum composed of three different aspects of compatibility. First, and as a bare minimum, the absence of serious aggression or injury demonstrates that pair-mates at least tolerate each other, and is a baseline feature of determining pair compatibility in most pairing programs at research facilities across the United States. Second, clear directionality in dominance signals between pair-mates indicates a certain and well established relationship (Pomerantz & Baker, 2017). Strongly compatible pair-mates will display these first two traits, as well as high levels of affiliative interaction, and score low on Tense as a pair when evaluated by staff with species-specific behavioral knowledge. We recommend, when possible, that behavioral management teams strive to match optimal pair-mates together but, when restricted, allow pair-mates to maintain consistency in their social interactions via continuous pair-housing, and use grates (if possible) when temporarily separating pairs overnight.

Research guiding the proper implementation of social housing is especially important for refining NHP welfare in the context of biomedical and basic research. Further research can improve biomedical and basic research project planning to mitigate physiological changes that may result from manipulations of the social environment, while maximizing the quality of life of the NHPs involved.

ACKNOWLEDGEMENTS

This research was supported in part by an NIH base grant awarded to the California National Primate Research Center (CNPRC; P51-OD011107-53). KC was supported by NIH grant R21 MH099361. The authors would like to thank the CNPRC Primate Services for providing study subjects and the CNPRC Primate Assay Laboratory Core for performing the urinary cortisol and creatinine assays. Additional thanks go to the CNPRC Population Behavioral Health Services and Animal Care staff for taking care of the monkeys and assisting with the implementation of our project. We would also like to thank two anonymous reviewers and the editor, Karen Bales, for comments and suggestions to improve the manuscript. DH, LC, BM, JV, and SS developed the research questions, hypotheses, and analytical methods. KC developed the pair adjective rating method. LC collected the behavioral data and conducted pair ratings. LC, AB, and DH collected the biological samples. LC prepared, processed, and stored the biological samples. BM provided laboratory supplies, equipment, and laboratory space for collecting, processing, and storing samples. DH performed and interpreted the statistical analyses. DH, LC, JV, and SW wrote and revised the manuscript with additional edits and revisions by BM, SS, KC, and AB.

FIGURE LEGENDS

Figure 1. Experimental design

Figure 2. Model 1 marginal plots for the main effects of: a) Dominant and b) Affiliation.

Figure 3. Model 1 margins plot of the interaction of Current Condition and Tense

REFERENCES

Abbott, D., Keverne, E., Bercovitch, F., Shively, C., Mendoza, S. P., Saltzman, W., . . . Garland, T. (2003). Are subordinates always stressed? A comparative analysis of rank differences in cortisol levels among primates. *Hormones and Behavior, 43*(1), 67-82.

Anderson, J. R. (1998). Sleep, sleeping sites, and sleep-related activities: Awakening to their significance. *American Journal of Primatology, 46*(1), 63-75. doi:10.1002/(SICI)1098-2345(1998)46:1<63::AID-AJP5>3.0.CO;2-T

Baker, K. C. (2016). Survey of 2014 behavioral management programs for laboratory primates in the United States. *American Journal of Primatology, 78*(7), 780-796. doi:10.1002/ajp.22543

Baker, K. C., Bloomsmith, M. A., Oettinger, B., Neu, K., Griffis, C., Schoof, V., & Maloney, M. (2012). Benefits of pair housing are consistent across a diverse population of rhesus macaques. *Applied Animal Behaviour Science, 137*(3), 148-156. doi:10.1016/j.applanim.2011.09.010

Baker, K. C., Bloomsmith, M. A., Oettinger, B., Neu, K., Griffis, C., & Schoof, V. A. (2014). Comparing options for pair housing rhesus macaques using behavioral welfare measures. *American Journal of Primatology, 76*(1), 30-42. doi:10.1002/ajp.22190

Baker, K. C., Coleman, K., Bloomsmith, M., McCowan, B., & Truelove, M. (2014). *Pairing rhesus macaques (Macaca mulatta): methodology and outcomes at four national primate research centers*. Paper presented at the American Journal of Primatology, Decatur, GA.

Baker, K. C., Crockett, C. M., Lee, G. H., Oettinger, B. C., Schoof, V., & Thom, J. P. (2012). Pair housing for female longtailed and rhesus macaques in the laboratory: behavior in

protected contact versus full contact. *Journal of Applied Animal Welfare Science, 15*(2), 126-143. doi:10.1080/10888705.2012.658330

Basile, B., Hampton, R., Chaudry, A., & Murray, E. (2007). Presence of a privacy divider increases proximity in pair-housed rhesus monkeys. *Animal Welfare, 16*(1), 37.

Bennett, B. T. (2016). Association of Primate Veterinarians 2014 Nonhuman Primate Housing Survey. *Journal of the American Association for Laboratory Animal Science, 55*(2), 172-174.

Burnham, K. P., & Anderson, D. R. (2001). Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research, 28*(2), 111-119. doi:http://dx.doi.org/10.1071/WR99107

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York: Springer.

Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology, 65*(1), 23-35.

Capitanio, J. P. (1999). Personality dimensions in adult male rhesus macaques: Prediction of behaviors across time and situation. *American Journal of Primatology, 47*(4), 299-320.

Capitanio, J. P., Blozis, S. A., Snarr, J., Steward, A., & McCowan, B. (2017). Do "birds of a feather flock together" or do "opposites attract"? Behavioral responses and temperament predict success in pairings of rhesus monkeys in a laboratory setting. *American Journal of Primatology*.

Capitanio, J. P., Kyes, R. C., & Fairbanks, L. A. (2006). Considerations in the selection and conditioning of Old World monkeys for laboratory research: Animals from domestic sources. *Institute for Laboratory Animal Research Journal, 47*(4), 294-306.

Capitanio, J. P., Mendoza, S. P., Lerche, N. W., & Mason, W. A. (1998). Social stress results in altered glucocorticoid regulation and shorter survival in simian acquired immune deficiency syndrome. *Proceedings of the National Academy of Sciences, 95*(8), 4714-4719.

Cavigelli, S. A., & Caruso, M. J. (2015). Sex, social status and physiological stress in primates: the importance of social and glucocorticoid dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1669). doi:10.1098/rstb.2014.0103

Coelho, A. M., Carey, K. D., & Shade, R. E. (1991). Assessing the effects of social environment on blood pressure and heart rates of baboons. *American Journal of Primatology, 23*(4), 257-267. doi:10.1002/ajp.1350230406

Crockett, C. M., Bowers, C. L., Bowden, D. M., & Sackett, G. P. (1994). Sex differences in compatibility of pair-housed adult longtailed macaques. *American Journal of Primatology, 32*(2), 73-94. doi:10.1002/ajp.1350320202

Crockford, C., Wittig, R. M., Whitten, P. L., Seyfarth, R. M., & Cheney, D. L. (2008). Social stressors and coping mechanisms in wild female baboons (*Papio hamadryas ursinus*). *Hormones and Behavior, 53*(1), 254-265. doi:10.1016/j.yhbeh.2007.10.007

Dettmer, A. M., Novak, M. A., Meyer, J. S., & Suomi, S. J. (2014). Population density-dependent hair cortisol concentrations in rhesus monkeys (*Macaca mulatta*). *Psychoneuroendocrinology, 42*(Supplement C), 59-67. doi:https://doi.org/10.1016/j.psyneuen.2014.01.002

Dettmer, E., & Fragaszy, D. (2000). Determining the value of social companionship to captive tufted capuchin monkeys (*Cebus apella*). *Journal of Applied Animal Welfare Science, 3*(4), 293-304. doi:10.1207/S15327604JAWS0304_2

Doyle, L. A., Baker, K. C., & Cox, L. D. (2008). Physiological and behavioral effects of social introduction on adult male rhesus macaques. *American Journal of Primatology, 70*(6), 542-550. doi:10.1002/ajp.20526

Dunbar, R. I. (1991). Functional significance of social grooming in primates. *Folia Primatologica, 57*(3), 121-131. doi:10.1159/000156574

Eaton, G. G., Kelley, S. T., Axthelm, M. K., Iliff-Sizemore, S. A., & Shiigi, S. M. (1994). Psychological well-being in paired adult female rhesus (*Macaca mulatta*). *American Journal of Primatology, 33*(2), 89-99. doi:10.1002/ajp.1350330204

Emery, N. J., Capitanio, J. P., Mason, W. A., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (*Macaca mulatta*). *Behavioral Neuroscience, 115*(3), 515. doi:10.I037//0735-7044.I15.3.515

Gilbert, M. H., & Baker, K. C. (2011). Social buffering in adult male rhesus macaques (*Macaca mulatta*): effects of stressful events in single vs. pair housing. *Journal of Medical Primatology, 40*(2), 71-78. doi:10.1111/j.1600-0684.2010.00447.x

Gottlieb, D. H., Maier, A., & Coleman, K. (2015). Evaluation of environmental and intrinsic factors that contribute to stereotypic behavior in captive rhesus macaques (*Macaca mulatta*). *Applied Animal Behaviour Science, 171*, 184-191. doi:10.1016/j.applanim.2015.08.005

Grueber, C., Nakagawa, S., Laws, R., & Jamieson, I. (2011). Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology, 24*(4), 699-711.

Gust, D. A., Gordon, T. P., Brodie, A. R., & McClure, H. M. (1994). Effect of a preferred companion in modulating stress in adult female rhesus monkeys. *Physiology & Behavior, 55*(4), 681-684. doi:10.1016/0031-9384(94)90044-2

Hamel, A. F., Lutz, C. K., Coleman, K., Worlein, J. M., Peterson, E. J., Rosenberg, K. L., . . . Meyer, J. S. (2017). Responses to the Human Intruder Test are related to hair cortisol phenotype and sex in rhesus macaques (Macaca mulatta). *American Journal of Primatology, 79*(1), e22526-n/a. doi:10.1002/ajp.22526

Hannibal, D., Bliss-Moreau, E., Vandeleest, J., McCowan, B., & Capitanio, J. P. (2017). Laboratory rhesus macaque social housing and social changes: Implications for research. *American Journal of Primatology, 79*(1), e22528.

Hardin, J. W., & Hilbe, J. (2007). *Generalized linear models and extensions* (2nd ed.). College Station, Tex.: Stata Press.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*, 77-89.

Hennessy, M. B., Kaiser, S., & Sachser, N. (2009). Social buffering of the stress response: diversity, mechanisms, and functions. *Frontiers in Neuroendocrinology, 30*(4), 470-482. doi:10.1016/j.yfrne.2009.06.001

Herman, J. P., Figueiredo, H., Mueller, N. K., Ulrich-Lai, Y., Ostrander, M. M., Choi, D. C., & Cullinan, W. E. (2003). Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo–pituitary–adrenocortical responsiveness. *Frontiers in Neuroendocrinology, 24*(3), 151-180. doi:https://doi.org/10.1016/j.yfrne.2003.07.001

Kikusui, T., Winslow, J. T., & Mori, Y. (2006). Social buffering: relief from stress and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1476), 2215-2228. doi:10.1098/rstb.2006.1941

Lilly, A. A., Mehlman, P. T., & Higley, J. D. (1999). Trait-like immunological and hematological measures in female rhesus across varied environmental conditions. *American Journal of Primatology, 48*(3), 197-223. doi:10.1002/(SICI)1098-2345(1999)48:3<197::AID-AJP3>3.0.CO;2-Y

Lutz, C. K., & Novak, M. A. (2005). Environmental enrichment for nonhuman primates: theory and application. *ILAR Journal, 46*(2), 178-191. doi:10.1093/ilar.46.2.178

Meagher, R. K. (2009). Observer ratings: validity and value as a tool for animal welfare research. *Applied Animal Behaviour Science, 119*(1), 1-14. doi:10.1016/j.applanim.2009.02.026

Mendoza, S., Capitanio, J. P., & Mason, W. (2001). Chronic social stress: Studies in nonhuman primates. In G. Moberg & J. Mench (Eds.), *Biology of Animal Stress: Basic Principles and Implications ofr Animal Welfare* (pp. 227-247). New York: CABI Publishing.

Mineka, S., Suomi, S. J., & DeLizio, R. (1981). Multiple separations in adolescent monkeys: An opponent-process interpretation. *Journal of Experimental Psychology: General, 110*(1), 56. doi:10.1037/0096-3445.110.1.56

Muller, M. N., & Wrangham, R. W. (2004). Dominance, cortisol and stress in wild chimpanzees (*Pan troglodytes schweinfurthii*). *Behavioral Ecology and Sociobiology, 55*(4), 332-340. doi:10.1007/s00265-003-0713-1

Novak, M. A., Hamel, A. F., Kelly, B. J., Dettmer, A. M., & Meyer, J. S. (2013). Stress, the HPA axis, and nonhuman primate well-being: a review. *Applied Animal Behaviour Science, 143*(2), 135-149.

Olsson, I. A. S., & Westlund, K. (2007). More than numbers matter: The effect of social factors on behaviour and welfare of laboratory rodents and non-human primates. *Applied Animal Behaviour Science, 103*(3–4), 229-254. doi:10.1016/j.applanim.2006.05.022

Pomerantz, O., & Baker, K. C. (2017). Higher levels of submissive behaviors at the onset of the pairing process of rhesus macaques (*Macaca mulatta*) are associated with lower risk of wounding following introduction. *American Journal of Primatology*. doi:10.1002/ajp.22671

Reeder, D. M., & Kramer, K. M. (2005). Stress in free-ranging mammals: Integrating physiology, ecology, and natural history. *Journal of Mammalogy, 86*(2), 225-235. doi:10.1644/BHE-003.1

Richter, S. H., Garner, J. P., Zipser, B., Lewejohann, L., Sachser, N., Touma, C., . . . Wurbel, H. (2011). Effect of population heterogenization on the reproducibility of mouse behavior: A multi-laboratory study. *Plos One, 6*(1). doi:10.1371/journal.pone.0016461

Roberts, S. J., & Platt, M. L. (2005). Effects of isosexual pair-housing on biomedical implants and study participation in male macaques. *Journal of the American Association for Laboratory Animal Science, 44*(5), 13-18.

Rommeck, I., Anderson, K., Heagerty, A., Cameron, A., & McCowan, B. (2009). Risk factors and remediation of self-injurious and self-abuse behavior in rhesus macaques. *Journal of Applied Animal Welfare Science, 12*(1), 61-72. doi:10.1080/10888700802536798

Rommeck, I., Capitanio, J., Strand, S. C., & McCowan, B. (2011). Early social experience affects behavioral and physiological responsiveness to stressful conditions in infant rhesus macaques (*Macaca mulatta*). *American Journal of Primatology, 73*(7), 692-701.

Sapolsky, R. M. (1992). Cortisol concentrations and the social significance of rank instability among wild baboons. *Psychoneuroendocrinology, 17*(6), 701-709. doi:10.1016/0306-4530(92)90029-7

Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews, 21*(1), 55-89. doi:10.1210/er.21.1.55

Schapiro, S. J., Bloomsmith, M. A., Kessel, A. L., & Shively, C. A. (1993). Effects of enrichment and housing on cortisol response in juvenile rhesus monkeys. *Applied Animal Behaviour Science, 37*(3), 251-263. doi:10.1016/0168-1591(93)90115-6

Schapiro, S. J., Nehete, P. N., Perlman, J. E., & Sastry, K. J. (2000). A comparison of cell-mediated immune responses in rhesus macaques housed singly, in pairs, or in groups. *Applied Animal Behaviour Science, 68*(1), 67-84. doi:10.1016/S0168-1591(00)00090-3

Shively, C. A. (1998). Social subordination stress, behavior, and central monoaminergic function in female cynomolgus monkeys. *Biological Psychiatry, 44*(9), 882-891.

Shutt, K., MacLarnon, A., Heistermann, M., & Semple, S. (2007). Grooming in Barbary macaques: better to give than to receive? *Biology Letters, 3*(3), 231-233. doi:10.1098/rsbl.2007.0052

Silk, J. B., Altmann, J., & Alberts, S. C. (2006). Social relationships among adult female baboons (*Papio cynocephalus*) I. Variation in the strength of social bonds. *Behavioral Ecology and Sociobiology, 61*(2), 183-195. doi:10.1007/s00265-006-0249-2

Smith, T. E., & French, J. A. (1997). Psychosocial stress and urinary cortisol excretion in marmoset monkeys. *Physiology & Behavior, 62*(2), 225-232. doi:10.1016/S0031-9384(97)00103-0

Symonds, M. R., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology, 65*(1), 13-21.

Tardif, S. D., Coleman, K., Hobbs, T. R., & Lutz, C. (2013). IACUC Review of Nonhuman Primate Research. *Institute for Laboratory Animal Research Journal, 54*(2), 234-245. doi:10.1093/ilar/ilt040

Truelove, M. A., Martin, A. L., Perlman, J. E., Wood, J. S., & Bloomsmith, M. A. (2017). Pair housing of macaques: A review of partner selection, introduction techniques, monitoring for compatibility, and methods for long-term maintenance of pairs. *American Journal of Primatology, 79*, e22485. doi:10.1002/ajp.22485

United States Department of Agriculture. (2013). *Animal Welfare Act and Animal Welfare Regulations*. http://awic.nal.usda.gov/government-and-professional-resources/federal-laws/animal-welfare-act: National Agricultural Library.

Vandeleest, J. J., Beisner, B. A., Hannibal, D. L., Nathman, A. C., Capitanio, J. P., Hsieh, F., . . . McCowan, B. (2016). Decoupling social status and status certainty effects on health in macaques: a network approach. *PeerJ, 4*, e2394. doi:10.7717/peerj.2394

Vandeleest, J. J., Blozis, S. A., Mendoza, S. P., & Capitanio, J. P. (2013). The effects of birth timing and ambient temperature on the hypothalamic–pituitary–adrenal axis in 3–4 month old rhesus monkeys. *Psychoneuroendocrinology, 38*(11), 2705-2712. doi:https://doi.org/10.1016/j.psyneuen.2013.06.029

Weed, J. L., Wagner, P. O., Byrum, R., Parrish, S., Knezevich, M., & Powell, D. A. (2003). Treatment of persistent self-injurious behavior in rhesus monkeys through socialization: a

preliminary report. *Journal of the American Association for Laboratory Animal Science, 42*(5), 21-23.

Wooddell, L. J., Hamel, A. F., Murphy, A. M., Byers, K. L., Kaburu, S. S. K., Meyer, J. S., . . . Dettmer, A. M. (2017). Relationships between affiliative social behavior and hair cortisol concentrations in semi-free ranging rhesus monkeys. *Psychoneuroendocrinology, 84*(Supplement C), 109-115. doi:https://doi.org/10.1016/j.psyneuen.2017.06.018

RESEARCH HIGHLIGHTS

- Overnight separations of tense pairs are associated with increased HPA-axis activity
- Continuous and compatible pair-housing are recommended
- When separating overnight, contact via bars or grates may improve welfare and alleviate unintended effects

Table 1. Variables included in model selection analyses, sorted in alphabetical order

Variable	Description
Abnormal	Proportion of focal intervals that included at least one instance of the following abnormal behaviors: regurgitate, urine/feces ingest, floating limb, self-strumming, leg lift, eye poke, suck (self or other), self-clasp, cheek biting, self-bite, threat-bite, self-hit, self-injurious behavior, hair pluck (self or other), hair ingest, pacing, swinging, flipping, twirling, rocking, bouncing, head twist, withdrawn
Affiliation	Proportion of focal intervals that included at least one instance of the following dyadic affiliative or prosocial behaviors: co-threat, recruit, recruit join, present ventrum/body, present rump, mount, mount solicited, anogenital exploration, play, huddle, reconcile, groom given, groom receive, mutual groom
Agonistic	Proportion of focal intervals that included at least one instance of the following agonistic behaviors: non-contact aggression (threat, lunge, cringe, display, redirect, response non-contact aggression), contact aggression (push, pull, slap, wrestle, grapple, chase, bite, hair pull, pin, response contact aggression), trauma (mild or severe),
Cohort	Whether the subject was in the first or second cohort
Current Condition	Current pairing condition (continuous or intermittent)
Dominant	Whether the animal is dominant to their pair-mate based on receiving the greatest proportion of status signaling behaviors (move away, turn away, silent bared teeth) displayed between them
Experimental Group	Began as intermittent and then experimentally changed to continuous (IC), or began as continuous and then experimentally changed to intermittent (CI)
Foraging enrichment	Whether the subjects received foraging enrichment prior to focal
Groom Given	Subject picks, scrapes, spreads, mouth picks and/or licks partner's hair or skin (not included in the same model with other groom variables or affiliation variable)

Groom Mutual	Subject and partner picks, scrape, spread, mouth pick and/or lick each other's hair or skin (not included in the same model with other groom variables or affiliation variable)
Groom Received	Partner picks, scrapes, spreads, mouth picks and/or licks subject's hair or skin (not included in the same model with other groom variables or affiliation variable)
Grooming	Proportion of focal intervals that included at least one instance of the following grooming behaviors: groom given, groom receive, mutual groom (not included in the same model with other groom variables or affiliation variable)
Inactive	Subject is not active for more than 5 seconds
Initial Pairing Condition	Subject's pairing condition at the beginning of the study
Initial Pairing Condition Time	Total time in months that the subject was living with current pair-mate in the initial housing condition before study
Menses	Subject's menstrual blood observed by husbandry staff.
Pair ID	The unique identification number for each pair to assess as a random effect
Pair Rating Anxious[a]	Score on pair rating measure "anxious" (seven-point scale): animals seek proximity when un-paired; pair is impatient during separation by vocalizing, manipulating pairing door, or being very eager to be re-paired
Pair Rating Friendly[a]	Score on pair rating measure "friendly" (seven-point scale): dyad enjoys the company of each other; both animals seek out social contact with partner; for example, playing, walking next to, or sitting with another monkey
Pair Rating Tense[a]	Score on pair rating measure "tense" (seven-point scale): pair is sociable to each other, but posture is rigid and not relaxed
Pair Rating Well-meshed[a]	Score on pair rating measure "well-meshed" (seven-point scale): animals are sensitive to each other in a non-anxious way

Project Phase	Current phase of the study (initial or experimental)
Related	Whether the subject's pair-mate is from the same matriline
Same Social Group	Whether the subject was reared in the same outdoor social group with its pair-mate
Status Signals Dominant	Subject approaches, sniffs the mouth of, or takes the resource (e.g., food or toy) of their pair-mate
Status Signals Subordinate	Subject moves away, turns away, averts eyes, freezes, or gives a silent bared teeth signal to their pair-mate
Study Week	The current week, out of 5, of the study
Total Pairing Time	Total time in months that the subject was living with current pair-mate before study
Total Time Indoors	Total time in months that the subject was living in indoor housing before study

[a]definition developed by K. Chun

Table 2. Best model set

	Model parameters	AICc	ΔAICc	L	*w*	Cumulative *w*	ER
M1	Y = Dominant + Affiliation + Experimental Group + Current Condition + Tense + Current Condition*Tense	3063.47	0.00	1.00	0.48	0.48	1.00
M2	Y = Dominant + Affiliation + Project Phase + Current Condition + Tense + Current Condition*Tense	3064.25	0.78	0.68	0.32	0.81	1.48
M3	Y = Dominant + Affiliation + Current Condition + Tense + Current Condition*Tense	3066.19	2.72	0.26	0.12	0.93	3.91
M4	Y = Dominant + Affiliation + Project Phase + Current Condition + Total Pairing Time + Inactive	3067.30	3.83	0.15	0.07	1.00	6.79

AICc: Corrected Akaike Information Criterion

ΔAICc: Difference in AICc value from that of M1

L: Model likelihood calculated from the formula $L(g_i|data) = \exp(-(1/2)\Delta AICc_i)$

w: The Akaike model weight ($L_i / \sum_{j=1}^{R} = L_j$. A measure of the strength of the evidence represented as a probability it is the best model.

ER: The evidence ratio, which is calculated by the weight of the best model divided by the weight of the given model.

Table 3. Variate weights for best model set

Variates	# of models	$w_+(j)$ [a]	Mean $w_+(j)$ [b]
Dominant	4	1.00	0.25
Affiliation	4	1.00	0.25
Current Condition	4	1.00	0.25
Tense	3	0.93	0.23
Project Phase	2	0.40	0.10
Experimental Group	1	0.48	0.12
Total Pairing Time	1	0.07	0.02
Inactive	1	0.07	0.02
Current Condition*Tense	3	0.93	0.23

[a] The sum of model weights that include the variate

[b] The proportion of the sum of the weights to the total number of models in the best model set

Table 4. Model 1 results

Variables Included in Model 1	β	β SE	exp(β)	exp(β) SE	exp(β) LBCI	exp(β) UBCI	P[a]	
Dominant	-0.497	0.165	0.608	0.101	0.440	0.841	0.003	**
Affiliation	-0.003	0.001	0.997	0.001	0.995	0.999	0.008	**
Experimental Group (IC)	-0.407	0.173	0.666	0.115	0.474	0.934	0.019	*
Current Condition (intermittent)	-0.604	0.183	0.547	0.100	0.382	0.782	0.001	**
Tense	-0.054	0.094	0.948	0.089	0.789	1.139	0.568	
Current Condition*Tense	0.262	0.063	1.299	0.082	1.147	1.471	<0.001	***

[a] Provided as probability information only, not as accept/reject criteria, which is not appropriate for an IT approach.

Significance denoted by: ***P<0.001, **P<0.01, *P<0.05

Cohort number	Experimental Group	Pairing change	Number of subjects		Study week number				
					Initial		Experimental		
			In study	In analysis	1	2	3	4	5
1	IC	Intermittent (I) → Continuous (C)	6	5	I	I	C	C	C
	CI	Continuous (C) → Intermittent (I)	6	6	C	C	I	I	I
2	IC	Intermittent (I) → Continuous (C)	6	4	I	I	C	C	C
	CI	Continuous (C) → Intermittent (I)	6	5	C	C	I	I	I

Figure 1. Experimental design

44x10mm (600 x 600 DPI)

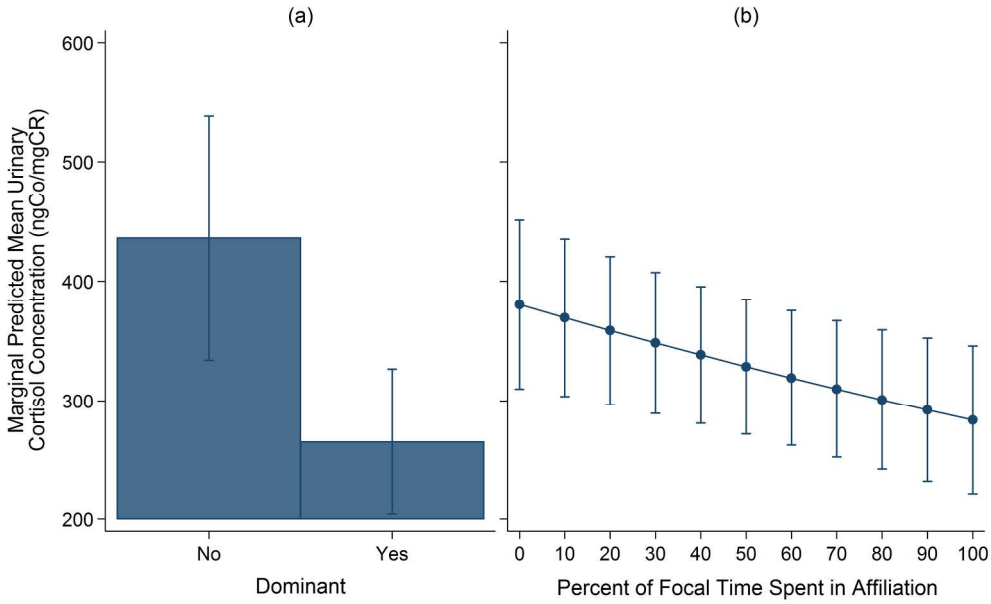

Figure 2. Model 1 marginal plots for the main effects of: a) Dominant and b) Affiliation.

111x69mm (600 x 600 DPI)

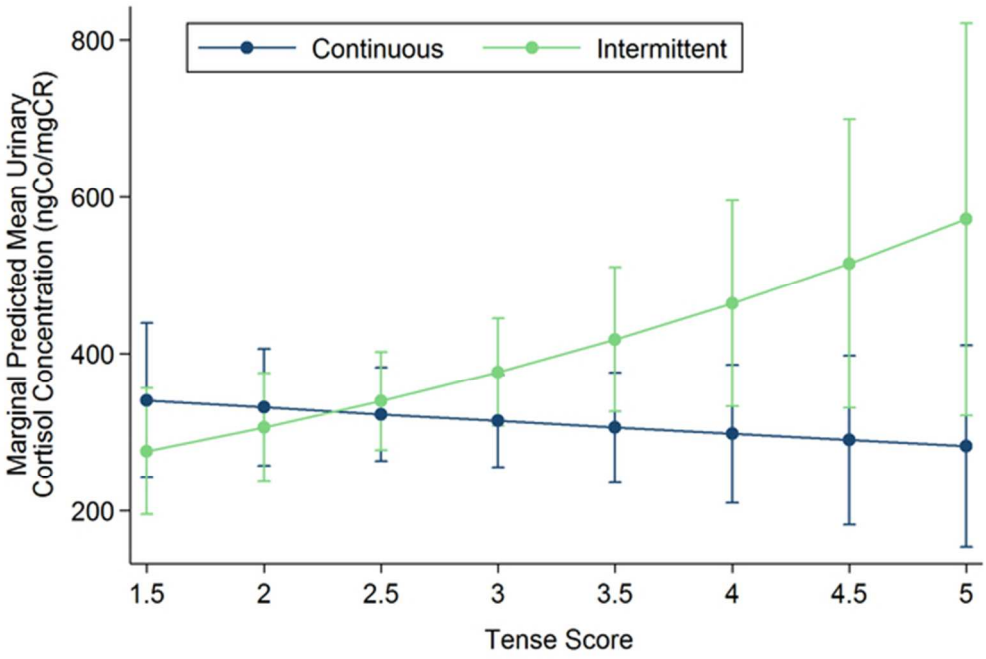

Figure 3. Model 1 margins plot of the interaction of Current Condition and Tense.

55x36mm (300 x 300 DPI)

Supplementary Table 1. Complete model set

Model	Retained	Reason for not Retaining	Independent Variables	Model Wald χ^2	Model P-value	AIC	AICc	ΔAIC	ΔAICc
M1	Yes		Dominant + Experimental Group + Affiliation + Current Condition + Tense + Current Condition*Tense	42.4	<0.0001	3062.84	3063.47	0.00	0.00
M2	Yes		Dominant + Project Phase + Affiliation + Current Condition + Tense + Current Condition*Tense	39.08	<0.0001	3063.62	3064.25	0.78	0.78
M3	Yes		Dominant + Affiliation + Current Condition + Tense + Current Condition*Tense	34.68	<0.0001	3065.71	3066.19	2.87	2.72
M4	Yes		Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Inactive	37.27	<0.0001	3066.67	3067.30	3.83	3.83
M5	Yes		Dominant + Experimental Group + Current Condition + Tense + Current Condition*Tense	34.44	<0.0001	3067.83	3068.31	4.99	4.84
M6	Yes		Dominant + Experimental Group + Affiliation + Total Pairing Time + Well-meshed + Current Condition + Current Condition*Well-meshed	41.46	<0.0001	3067.80	3068.59	4.96	5.12
M7	Yes		Dominant + Project Phase + Current Condition + Tense + Current Condition*Tense	30.46	<0.0001	3069.13	3069.61	6.29	6.15
M8	Yes		Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Well-meshed	36.45	<0.0001	3069.20	3069.83	6.36	6.36
M9	Yes		Dominant + Experimental Group + Total Pairing Time + Current Condition + Tense + Current Condition*Tense	35.22	<0.0001	3069.39	3070.02	6.55	6.55
	Yes		Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition	30.83	<0.0001	3070.34	3070.82	7.50	7.36
	Yes		Dominant + Current Condition + Tense + Current	26.47	<0.0001	3070.96	3071.33	8.13	7.86

		Condition*Tense						
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Tense	32.67	<0.0001	3071.25	3071.88	8.41	8.41
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Foraging enrichment	31.44	<0.0001	3071.67	3072.29	8.83	8.83
Yes		Dominant + Project Phase + Affiliation + Total Pairing Time	26.49	<0.0001	3071.97	3072.33	9.13	8.86
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Agonistic	31.15	<0.0001	3071.93	3072.56	9.09	9.09
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Friendly	31.45	<0.0001	3071.97	3072.60	9.13	9.13
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Anxious	30.89	<0.0001	3072.29	3072.92	9.45	9.45
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Abnormal	30.83	<0.0001	3072.32	3072.94	9.48	9.48
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Experimental Group	28.94	<0.0001	3072.50	3072.98	9.66	9.52
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Tense	28.78	<0.0001	3072.59	3073.07	9.75	9.61
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Cohort	28.6	<0.0001	3072.73	3073.22	9.89	9.75
Yes		Dominant + Project Phase + Affiliation	21.6	0.0001	3073.30	3073.55	10.46	10.09
No	Not parsimonious	Dominant + Project Phase + Affiliation + Total Pairing Time + Current Condition + Friendly + Current Condition*Friendly	32.76	<0.0001	3072.90	3073.69	10.07	10.22
No	Not	Dominant + Project Phase + Affiliation + Total Time	24.29	0.0001	3073.39	3073.75	10.55	10.29

	parsimonious	Indoors						
Yes		Current Condition + Tense + Current Condition*Tense	20.43	<0.0001	3074.01	3074.26	11.17	10.80
No		Dominant + Project Phase + Affiliation + Cohort	21.76	0.002	3075.18	3075.55	12.34	12.08
No	Not parsimonious	Dominant + Experimental Group + Affiliation + Total Pairing Time + Current Condition + Well-meshed	29.48	<0.0001	3076.90	3077.52	14.06	14.06
Yes		Dominant + Project Phase	14.64	0.0007	3077.67	3077.84	14.83	14.37
Yes		Dominant + Experimental Group + Affiliation	19.3	0.0002	3077.62	3077.87	14.78	14.41
No	Not parsimonious	Dominant + Project Phase + From Same Social Group	17.18	0.0007	3077.83	3078.08	14.99	14.62
No	Not parsimonious	Dominant + Project Phase + Total Time Indoors	16.87	0.0008	3078.03	3078.29	15.19	14.82
No	Not parsimonious	Dominant + Experimental Group + Affiliation + Total Pairing Time + Current Condition	23.3	0.0003	3078.14	3078.62	15.30	15.16
No	Not parsimonious	Dominant + Experimental Group + Groom Given	17.52	0.0006	3078.75	3079.01	15.91	15.54
No	Not parsimonious	Dominant + Experimental Group + Affiliation + Total Pairing Time + Current Condition + Tense	25.89	0.0002	3078.72	3079.35	15.88	15.88
No		Dominant + Experimental Group + Affiliation + Total Pairing Time + Current Condition + Inactive	24.65	0.0004	3078.83	3079.35	15.99	15.88
No	Not parsimonious	Dominant + Project Phase + Menses	14.95	0.0019	3079.42	3079.68	16.58	16.21
No	Not parsimonious	Dominant + Experimental Group + Grooming	17.22	0.0006	3079.48	3079.74	16.64	16.28
No	Not parsimonious	Dominant + Experimental Group + Affiliation + Cohort	19.69	0.0006	3079.41	3079.77	16.57	16.30

No	Not parsimonious	Dominant + Experimental Group + Affiliation + Total Pairing Time	19.55	0.0006	3079.48	3079.84	16.64	16.37
No	Not parsimonious	Dominant + Project Phase + Cohort	14.69	0.0021	3079.63	3079.89	16.79	16.42
Yes		Dominant + Affiliation	12.88	0.0016	3079.77	3079.94	16.93	16.48
No	Not parsimonious	Dominant + Experimental Group + Affiliation + Total Pairing Time + Tense	22.76	0.0004	3079.69	3080.17	16.85	16.71
Yes		Dominant + Groom Given	11.86	0.0027	3080.61	3080.78	17.77	17.31
Yes		Project Phase	8.52	0.0035	3081.02	3081.13	18.18	17.66
No	Not parsimonious	Dominant + Affiliation + Menses	13.96	0.003	3080.89	3081.15	18.05	17.68
No	Not parsimonious	Dominant + Affiliation + Agonistic	13.41	0.0038	3081.23	3081.28	18.39	17.81
Yes		Dominant + Grooming	11.04	0.004	3081.53	3081.70	18.69	18.24
Yes		Dominant + Experimental Group	12.49	0.0019	3081.83	3082.00	18.99	18.54
No	Not parsimonious	Project Phase + Menses	8.66	0.0132	3082.89	3083.06	20.05	19.59
No	Not parsimonious	Dominant + Experimental Group + Agonistic	12.91	0.0048	3083.27	3083.53	20.43	20.06
No	Not parsimonious	Dominant + Experimental Group + Menses	13.15	0.0043	3083.32	3083.58	20.48	20.11
Yes		Affiliation (all dyadic affiliative behavior and recruit and cothreat behaviors)	6.02	0.0141	3083.54	3083.64	20.70	20.18
No	Not parsimonious	Dominant + Experimental Group + From Same Social Group	12.72	0.0053	3083.70	3083.70	20.86	20.23

No	Not parsimonious	Dominant + Experimental Group + Total Pairing Time	12.74	0.0052	3083.68	3083.93	20.84	20.47
No	Not parsimonious	Dominant + Experimental Group + Cohort	12.64	0.0055	3083.75	3084.00	20.91	20.54
Yes		Dominant	6.33	0.0119	3083.97	3084.07	21.13	20.60
Yes		Groom Given	5.66	0.0174	3084.01	3084.11	21.17	20.65
Yes		Study Week (not used further because Project Phase performed better and is essentially a coarser recode of this variable)	11.79	0.019	3083.76	3084.12	20.92	20.65
No	Not parsimonious	Dominant + Current Condition	8.44	0.0147	3083.98	3084.15	21.14	20.69
No	Not parsimonious	Dominant + From Same Social Group	8.64	0.0133	3084.31	3084.32	21.47	20.86
No	Not parsimonious	Affiliation + Menses	6.74	0.0344	3084.88	3085.05	22.04	21.58
Yes		Grooming	4.39	0.0361	3085.15	3085.25	22.31	21.78
No	Not parsimonious	Dominant + Agonistic	7.06	0.0293	3085.21	3085.38	22.37	21.91
No	Not parsimonious	Affiliation + Cohort	6.33	0.0422	3085.24	3085.41	22.40	21.94
No	Not parsimonious	Dominant + Menses	7.02	0.0298	3085.38	3085.55	22.54	22.08
Yes		Related	7	0.03036	3085.47	3085.64	22.63	22.17
No	Not parsimonious	Dominant + Tense	6.85	0.0325	3085.57	3085.74	22.73	22.27

No	Not parsimonious	Dominant + Current Condition + Tense	8.85	0.0313	3085.68	3085.93	22.84	22.47
No	Not parsimonious	Dominant + Cohort	6.42	0.0404	3085.90	3086.07	23.06	22.61
No	Model χ^2 too low	Total Time Indoors	3.04	0.0811	3086.63	3086.73	23.79	23.27
No	Model χ^2 too low	Total Pairing Time	2.37	0.124	3087.22	3087.22	24.38	23.76
No	Model χ^2 too low	Experimental Group	2.31	0.1282	3087.27	3087.27	24.43	23.80
No	Model χ^2 too low	Current Condition	2.05	0.1519	3087.41	3087.41	24.57	23.94
Yes		AnimalID random effects only (empty model)	.	.	3087.45	3087.50	24.62	24.04
No	Model χ^2 too low	Foraging enrichment	1.59	0.2077	3087.85	3087.85	25.01	24.38
No	Model χ^2 too low	Mutual Groom	1.43	0.2312	3088.12	3088.12	25.28	24.66
No	Model χ^2 too low	From Same Social Group	1.28	0.2581	3088.12	3088.12		
No	Model χ^2 too low	Initial Pairing Condition (Not used further because similar to Total Pairing Time)	1.08	0.2995	3088.41	3088.41	25.57	24.94
No	Model χ^2 too low	Well-meshed	0.95	0.3308	3088.53	3088.53	25.69	25.06
No	Model χ^2 too low	Agonistic	0.5	0.4783	3088.94	3088.94	26.10	25.47
No	Model χ^2 too	Tense	0.42	0.5149	3089.03	3089.03	26.20	25.57

	low							
No	Model χ^2 too low	Experimental Group + Menses	2.68	0.262	3088.92	3089.09	26.08	25.62
No	Animal ID alone is better	PairID + AnimalID nested random effects only (empty model)	.	.	3088.99	3089.10	26.16	25.63
No	Model χ^2 too low	Menses	0.43	0.5144	3089.04	3089.14	26.20	25.67
No	Model χ^2 too low	Groom Received (not used further because of other groom variables)	0.3	0.5808	3089.15	3089.15	26.31	25.69
No	Model χ^2 too low	Experimental Group + Cohort	2.55	0.2796	3089.06	3089.23	26.22	25.77
No	Model χ^2 too low	Current Condition + Tense	2.4	0.3016	3089.08	3089.25	26.24	25.78
No	Model χ^2 too low	Inactive	0.17	0.6816	3089.29	3089.29	26.45	25.82
No	Model χ^2 too low	Abnormal	0.08	0.7788	3089.38	3089.38	26.54	25.91
No	Model χ^2 too low	Cohort	0.18	0.675	3089.28	3089.38	26.44	25.92
No	Model χ^2 too low	Total Submissive Behaviors	0.06	0.8048	3089.39	3089.39	26.55	25.93
No	Model χ^2 too low	Friendly	0.05	0.8195	3089.40	3089.40	26.56	25.94
No	Model χ^2 too low	Total Dominant Behaviors	0.11	0.7431	3089.35	3089.45	26.51	25.99
No	Model χ^2 too	Anxious	0	0.9937	3089.45	3089.45	26.62	25.99

	low								
No	Animal ID alone is better	PairID random effects only (empty model)	.	.	3142.87	3142.92	80.03	79.46

Supplementary Table 2. Candidate model set

	Model parameters	AICc[a]	ΔAICc[b]	L[c]	w[d]	Cumulative w[e]	ER[f]
M1	Y = Dominant + Affiliation + Experimental Group + Current Condition + Tense + Current Condition*Tense	3063.47	0.00	1.00	0.42	0.42	
M2	Y = Dominant + Affiliation + Project Phase + Current Condition + Tense + Current Condition*Tense	3064.25	0.78	0.68	0.28	0.71	1.48
M3	Y = Dominant + Affiliation + Current Condition + Tense + Current Condition*Tense	3066.19	2.72	0.26	0.11	0.81	3.91
M4	Y = Dominant + Affiliation + Project Phase + Current Condition + Total Pairing Time + Inactive	3067.30	3.83	0.15	0.06	0.88	6.79
M5	Y = Dominant + Experimental Group + Current Condition + Tense + Current Condition*Tense	3068.31	4.84	0.09	0.04	0.91	11.27
M6	Y = Dominant + Affiliation + Experimental Group + Total Pairing Time + Current Condition + Well-meshed + Current Condition*Well-meshed	3068.59	5.12	0.08	0.03	0.95	12.95
M7	Y = Dominant + Project Phase + Current Condition + Tense + Current Condition*Tense	3069.61	6.15	0.05	0.02	0.97	21.64
M8	Y = Dominant + Affiliation + Project Phase + Total Pairing Time+ Current Condition + Well-meshed	3069.83	6.36	0.04	0.02	0.98	24.06
M9	Y = Dominant + Experimental Group + Total Pairing Time + Current Condition + Tense + Current Condition*Tense	3070.02	6.55	0.04	0.02	1.00	26.50

[a]AICc: Corrected Akaike Information Criterion
[b]ΔAICc: Difference in AICc value from that of M1
[c]L: Model likelihood calculated from the formula $L(g_i|data) = \exp(-(1/2)\Delta AICc_i)$
[d]w: The Akaike model weight ($L_i / \sum_{j=1}^{R} = L_j$. A measure of the strength of the evidence for that model, represented as a probability.
[e]ER: The evidence ratio, which is calculated by the weight of the best model divided by the weight of the given model